Research Article | Vol. 14, No. 8 / 1 Aug 2023 / *Biomedical Optics Express* 4217

Biomedical Optics EXPRESS

Test-retest reliability of fNIRS in resting-state cortical activity and brain network assessment in stroke patients

GONGCHENG XU,[1,2] CONGCONG HUO,[1] JIAHUI YIN,[3] YANBIAO ZHONG,[4] GUOYU SUN,[5] YUBO FAN,[1,6] DAIFA WANG,[1,8] AND ZENGYONG LI[2,7,9]

[1] *Key Laboratory for Biomechanics and Mechanobiology of Ministry of Education, Beijing Advanced Innovation Center for Biomedical Engineering, School of Biological Science and Medical Engineering, Beihang University, Beijing, China*
[2] *Beijing Key Laboratory of Rehabilitation Technical Aids for Old-Age Disability, National Research Center for Rehabilitation Technical Aids, Beijing, China*
[3] *School of Athletic Performance, Shanghai University of Sport, Shanghai, China*
[4] *Department of Rehabilitation Medicine, First Affiliated Hospital of Gannan Medical University, Ganzhou, China*
[5] *Changsha Medical University, Changsha, China*
[6] *School of Engineering Medicine, Beihang University, Beijing, China*
[7] *Key Laboratory of Neuro-functional Information and Rehabilitation Engineering of the Ministry of Civil Affairs, Beijing, China*
[8] *daifa.wang@buaa.edu.cn*
[9] *lizengyong@nrcrta.cn*

Abstract: Resting-state functional near infrared spectroscopy (fNIRS) scanning has attracted considerable attention in stroke rehabilitation research in recent years. The aim of this study was to quantify the reliability of fNIRS in cortical activity intensity and brain network metrics among resting-state stroke patients, and to comprehensively evaluate the effects of frequency selection, scanning duration, analysis and preprocessing strategies on test-retest reliability. Nineteen patients with stroke underwent two resting fNIRS scanning sessions with an interval of 24 hours. The haemoglobin signals were preprocessed by principal component analysis, common average reference and haemodynamic modality separation (HMS) algorithm respectively. The cortical activity, functional connectivity level, local network metrics (degree, betweenness and local efficiency) and global network metrics were calculated at 25 frequency scales $\times$ 16 time windows. The test-retest reliability of each fNIRS metric was quantified by the intraclass correlation coefficient. The results show that (1) the high-frequency band has higher ICC values than the low-frequency band, and the fNIRS metric is more reliable than at the individual channel level when averaged within the brain region channel, (2) the ICC values of the low-frequency band above the 4-minute scan time are generally higher than 0.5, the local efficiency and global network metrics reach high and excellent reliability levels after 4 min ($0.5 < \text{ICC} < 0.9$), with moderate or even poor reliability for degree and betweenness ($\text{ICC} < 0.5$), (3) HMS algorithm performs best in improving the low-frequency band ICC values. The results indicate that a scanning duration of more than 4 minutes can lead to high reliability of most fNIRS metrics when assessing low-frequency resting brain function in stroke patients. It is recommended to use the global correction method of HMS, and the reporting of degree, betweenness and single channel level should be performed with caution. This paper provides the first comprehensive reference for resting-state experimental design and analysis strategies for fNIRS in stroke rehabilitation.

#491610 https://doi.org/10.1364/BOE.491610
Journal © 2023 Received 28 Mar 2023; revised 24 Jun 2023; accepted 5 Jul 2023; published 24 Jul 2023

Biomedical Optics EXPRESS

1. Introduction

In the last 30 years, functional near-infrared spectroscopy has gained widespread use in neuroscience research thanks to continuous advances in technical methods such as light source-detector arrangements, simulation of light propagation in the head [1,2], image reconstruction and data analysis [3–5]. Because of its advantages of safety, noninvasiveness, ease of mobility, and resistance to motion interference, an increasing number of researchers in the field of rehabilitation have chosen fNIRS as a research tool to explore the neuroimaging of patients, especially in the rehabilitation process of stroke patients [6–9]. Stroke is an injury to the central nervous system caused by the rupture or blockage of cerebral blood vessels, and its recovery is a complex intertwined process of motor, cognitive, and speech dysfunction and neurological recovery [10,11]. fNIRS can be used with specific dysfunctional task paradigms (upper and lower limb movements [12], gait [13], postural balance [14], constraint-induced movements [15], etc.) to provide real-time dynamic detection of brain functional activity while performing the task, evaluate the brain activation pattern, and reflect neural remodelling through longitudinal follow-up studies in stroke patients.

Sufficiently high test-retest consistency is critical in longitudinal studies [16,17], and to account for changes in fNIRS signals over the course of recovery, it is important to confirm the reliability of the technique as applied to neuroimaging in stroke. However, to date, there has not been a study of the test-retest reliability of fNIRS in stroke patients. This may be due to the difficulty in ensuring consistent engagement and cooperation in stroke patients performing repeated measurement tasks before and after. The simple and stable process of resting-state measurements provides an effective solution to this problem. Resting-state fNIRS studies have reported deficits in frontoparietal cortical functional connectivity in stroke patients compared to healthy brains [18], and longitudinal data have shown improved connectivity between primary motor, somatosensory, and premotor areas in the affected hemisphere [7]. Furthermore, although stroke lesions are usually concentrated around specific vascular regions, neuroimaging and statistical studies have confirmed that stroke has a global impact on the entire brain and its network properties and is therefore considered a network disease [11,19]. The network-based approach to research can provide greater insight into resolving stroke-related neurological deficits and recovery [20,21]. If network metrics can demonstrate strong test-retest reliability, they have the potential to be used as biomarkers for improving prognostic ability and therapeutic interventions in further poststroke studies in both acute and chronic phases. Therefore, the main purpose of this study is to quantify the test-retest reliability of fNIRS in cortical activity intensity and brain network assessment through resting-state measurement in patients with stroke.

A number of studies have reported the test-retest reliability of fNIRS outside the field of stroke rehabilitation, using the intraclass correlation coefficient (ICC) as the measure of test-retest reliability [22–24]. Most studies analysed the number of activated channels, activation level and activation imaging consistency of two or three fNIRS measurements in the time range of 1 day to several months based on specific task stimulation. Only a few fNIRS studies on graph theory parameters have confirmed the good reliability of binarization and weight networks in resting-state measurements of healthy adults [25,26]. One of them performed a comparison of ICC between two preprocessing methods, i.e., filtering at 0.01-0.08 Hz (commonly considered the frequency band of neural activity) and independent component analysis (ICA); there were no significant differences [25]. The results of another study on an auditory task confirmed that the haemodynamic modality separation (HMS) preprocessing algorithm significantly improved the reliability of fNIRS measurements [27]. It is difficult to summarize the effect of different preprocessing methods on fNIRS test-retest reliability based on previously published studies.

In fact, there is no consensus regarding the signal processing methods of fNIRS imaging [28,29]. The presence of interfering components such as slow hemodynamic signal, systemic physiological noise, scalp interference [30–32] and motion artifacts in the fNIRS signal [33,34]

Biomedical Optics EXPRESS

may lead to a high false-incidence rate or high false-negative rate in inter-channel correlation analysis [35]. The aim of signal preprocessing is to remove haemodynamic changes unrelated to neurovascular coupling, and it is considered that the most promising method to remove this interference is the combination of short-distance channel measurement and a regression algorithm [6,28]. However, popular commercial fNIRS systems usually do not have the hardware facilities of short-distance channels. Under the condition that it is impossible to arrange short-distance channels, the commonly used preprocessing methods to remove scalp and global blood flow noise to some extent are principal component analysis (PCA) [36], common average reference (CAR) [37], HMS [38] and ICA [39] algorithms. Although reliability is not the only parameter to consider when choosing data processing and analysis strategies, it can be used as an important tool to help select processing methods. Therefore, to improve the generalization of the findings, another aim of this study is to explore the impact of different preprocessing methods on the test-retest reliability of fNIRS. Here, we restrict our focus to algorithms for automatic processing (i.e., PCA, CAR, and HMS). Because the ICA algorithm requires visual exclusion of uninteresting components in the intermediate process, thereby introducing human interference, it is not included in the research approach of this paper.

In addition, the measurement time of neuroimaging and the choice of analysis frequency band have been confirmed to have a significant effect on ICC in functional magnetic resonance imaging (fMRI) studies. Zuo et al. reported that regional homogeneity showed a reliability level of more than 50% in a time series of more than 4 minutes [40]. The reliability of standard bandpass-filtered data (approximately 0.01-0.1 Hz) is lower than that of other wider band-filtered data [17]. However, no fNIRS test-retest reliability studies have focused on these two important variables. A convenient method to study the signal characteristics in the time-frequency domain is the continuous wavelet transform (CWT) [41,42]. Through the selection of an appropriate mother wavelet function and central frequency, CWT can transform time series from the time domain to the time-frequency domain and ensure good time-frequency resolution. This provides a suitable analytical tool for this paper to explore the impact of measurement duration and band selection on the reliability of fNIRS for stroke rehabilitation.

This paper is based on the dataset of two resting-state fNIRS scanning sessions of patients with subacute stroke (within 3 months after stroke). Considering the obvious changes in brain function in patients with subacute stroke on a week-based time scale [43,44], the interval between the two sessions was only 24 hours. On this basis, the collected fNIRS signals were processed by PCA, CAR and HMS. The intensity of brain activity and the characteristics of the brain network on multiple time and frequency scales were calculated by wavelet transform and graph theory algorithms. Then, the test-retest reliability of the corresponding preprocessing method and time-frequency scale characteristics between the two sessions was analysed. To our knowledge, this is the first test-retest reliability study on fNIRS application in resting-state cortical activity and brain networks in patients with stroke. We aim to conduct a comprehensive reliability assessment and to provide valuable measurement and analysis references for the application of fNIRS in the stroke rehabilitation field.

2. Methods

This article follows the fNIRS studies guideline [45] to report on the research.

2.1. Subjects and fNIRS data acquisition

A total of 19 stroke subjects were recruited. The admission criteria were as follows: (1) within 2 weeks to 3 months after stroke and (2) unilateral lesions. The exclusion criteria were as follows: (1) severe dysfunction or failure of the heart, liver, lung, kidney and other important organs; (2) obvious cognitive impairment, which makes it impossible for participants to understand and cooperate with the explanation; and (3) continuous deterioration of the disease, new infarction or

Biomedical Optics EXPRESS

secondary bleeding. All subjects provided written informed consent before participating, and the study was approved by the Human Ethics Committee of the National Research Center for Rehabilitation Technical Aids. Table 1 shows the clinical details of the subjects.

Table 1. Clinical details.

Subjects	Sex	Age	BMI	Type of stroke	Affected side	Time post-stroke (day)	Site of lesion	BI	NIHSS	Fugl-Meyer		DPF	
										UE	LE	760 nm	850 nm
1	M	32	33.0	H	R	85	L, BG	50	7	17	25	6.35	5.29
2	M	49	29.4	I	L	45	R, BS	85	4	31	32	6.82	5.76
3	M	57	22.5	I	R	33	L, BG, TR	90	3	31	26	7.03	5.97
4	F	63	23.3	H	L	53	R, BS	40	9	4	14	7.18	6.12
5	F	61	26.0	I	R	81	L, BG, CR	70	5	30	32	7.13	6.07
6	M	42	24.6	H	L	26	R, BG	75	4	31	34	6.63	5.57
7	M	22	21.6	I	L	90	R, BG	85	3	56	34	6.06	5.00
8	M	61	28.7	H	L	74	R, FL, PL	35	9	6	16	7.13	6.07
9	M	60	26.0	H	L	94	R, BG, TR, CR	80	3	34	30	7.11	6.04
10	M	67	24.6	I	L	76	R, FL, TL, INS, BG	75	7	20	16	7.29	6.22
11	M	30	32.6	H	R	37	L, BG,	95	2	60	34	6.30	5.23
12	M	49	20.1	I	R	53	L, PO	90	2	44	34	6.82	5.76
13	F	52	24.8	I	L	84	R, BS, CER	50	6	20	22	6.90	5.84
14	F	69	23.4	I	L	64	R, BS	20	11	4	0	7.34	6.27
15	M	26	23.0	I	R	34	L, BG, CR	85	3	30	29	6.18	5.12
16	M	49	28.9	H	R	39	L, BG, TR	55	7	20	26	6.82	5.76
17[a]	M	40	25.6	H	L	79	R, BG	90	4	31	25	6.58	5.51
18[a]	M	30	31.2	H	L	83	R, FL	55	5	25	22	6.30	5.23
19[a]	M	47	26.0	I	L	7	R, BG	10	8	19	24	6.77	5.70

BG, basal ganglia; BI, the Barthel index of ADL; BMI, body mass index; BS, brain stem; CER, cerebellum; CR, corona radiata; DPF, the differential pathlength factor [46]; F, female; FL, frontal lobe; L, left; LE, Fugl–Meyer Assessment-lower extremities; H, hemorrhagic; I, ischemic; INS, insula; L, left; M, male; NIHSS, NIH stroke scale; PL, parietal lobe; PO, pons; R, right; TL, temporal lobe; TR, thalamic region; UE, Fugl–Meyer Assessment-upper extremities.
[a]means excluded.

The continuous wave fNIRS instrument NirSmart (Danyang Huichuang Medical Equipment Co., Ltd., China) was applied to acquire resting state fNIRS data twice with a duration of 9 minutes and an interval of 24 hours, the scan time of each subject was fixed at 5 pm. The acquisition time was set to 9 minutes to include at least five low frequency periods (0.01 Hz) to ensure the effectiveness of subsequent phase-related analysis. During the two scanning sessions, the subjects were asked to sit quietly on a stool with armrests facing the same direction in the same room, staying relaxed and ensuring that they were awake to avoid falling asleep. On the side of the subjects, the researchers made sure that the subjects' eyes were fixed on the white wall in front of them, and were reminded with a recorded slight taping sound when they were sleepy or distracted. Sixteen light sources and 7 probes constituted 24 measurement channels of 3 cm, which were arranged in the left and right prefrontal cortex (LPFC/RPFC) and the left and right motor cortex (LMC/RMC), as shown in Fig. 1(a). When wearing the optode cap, it was ensured that the Cz point of the electrode cap coincided with the Cz point of the scalp surface measurement, and the consistency of positioning at the individual level was further ensured by referring to the photograph taken at the first session. The sampling frequency was set to 10 Hz using two wavelengths (760 nm and 850 nm) for scanning.

Biomedical Optics EXPRESS

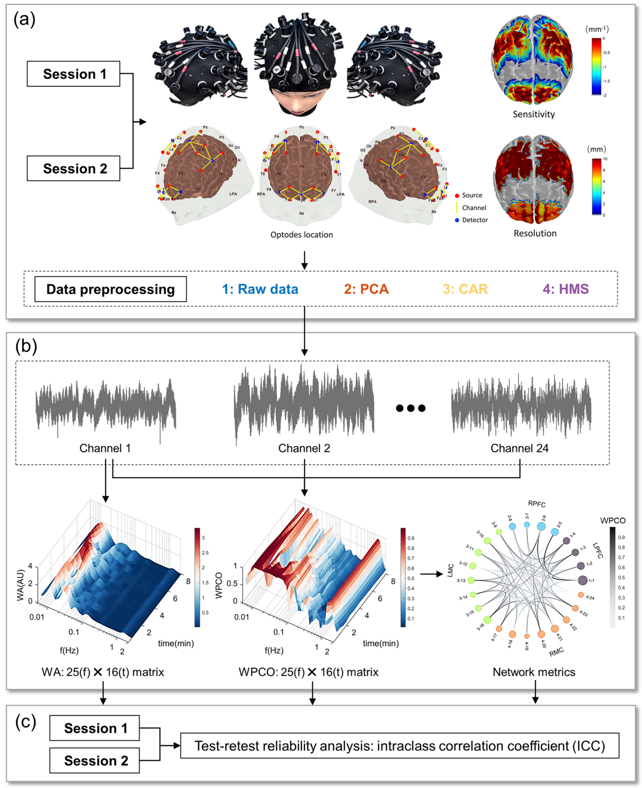

Fig. 1. Flow chart of cortical activity and brain network reliability analysis at multiple preprocessing modalities and multiple time-frequency scales. (a) The light sources and probes arrangement corresponding standardized 10-20 location and different preprocessing steps for the two scanning sessions. Four channels were arranged around Fp1 and Fp2 in the LPFC and RPFC, respectively. Eight channels were arranged around C3 and C4 in the LMC and RMC, respectively. The location, sensitivity and resolution maps were created using AtlasViewer [47]. (b) WA was calculated from each channel signal, WPCO was calculated from every pairwise channel, network matrix based on WPCO value was constructed with 50% sparsity and network metrics were calculated. The grayscale value of the network map indicates the magnitude of the WPCO value. (c) Calculation of intraclass correlation coefficients for fNIRS metrics obtained by two scanning sessions.

2.2. Data preprocessing

It should be noted in advance that for all subjects, the affected hemisphere was indicated as the right side. The cortical haemoglobin concentration signals parsed by resting fNIRS data are usually disturbed by motion artefacts, breathing, heartbeat and scalp noise [48,49]. A common practice to eliminate the effects of respiration and heartbeat is by low-pass filtering below 0.2 Hz, while this study will use the CWT algorithm to resolve the signal features corresponding to 25 frequency scales in the 0.01-2 Hz band, thus omitting the low-pass filtering step. The overall preprocessing process is performed in MATLAB software as follows:

(1) Remove the first minute data to ensure data stability. For a single channel of light intensity signal, if the intensity >1000 or intensity <0.5 or the mean/standard deviation of the intensity time series <2, it was considered a noisy channel [50]. The cortical activity and brain network metrics calculated from the noisy channels will be excluded from the subsequent analysis. If the

number of noisy channels exceeded 25% of the total number of channels ($24 \times 25\% = 6$), the subject was considered to have low signal quality and needed to be excluded. A total of 3 subjects were excluded from this study, leaving 16 subjects to continue with the follow-up analysis.(2) Motion artefact interference was removed by a time derivative distribution repair algorithm [51]. Signal spikes were removed by the Hample filtering algorithm with filter parameters of $k = 50$ and $nsigma = 4$. The 8th order bandpass Butterworth filter of 0.01-2 Hz was used to remove low frequency drift and high frequency instrument noise.(3) The hmrIntensity2OD and hmrOD2Conc functions in the Homer2 toolbox were used to convert light intensity signals into oxygenated haemoglobin (HbO_2) and deoxyhaemoglobin (HbR) signals [52].(4) Three different algorithms for purifying the haemodynamic changes related to neurovascular coupling were performed on HbO_2 and HbR signals:

a. PCA

The working principle of PCA is to eliminate several components with the greatest contribution to variance, which correspond to the first few spatial covariances of fNIRS data and are considered to represent the global impact of all channels [28]. PCA typically performs better in multichannel fNIRS measurements across brain regions, similar to the design of this study. The number of PCA components removed in this study was set to explain 80% of the variance [29].

b. CAR

Based on the assumption that global noise affects all fNIRS channels, the CAR algorithm subtracted the mean value of all channel time series from the time series of each channel to reduce global interference [53]. Similar to PCA, the CAR algorithm is more effective in multiple brain functional area measurement datasets than local measurements. We excluded channels with low signal-to-noise ratios when performing the procedure of averaging over all channels.

c. HMS

The HMS algorithm is based on the fact that there is a negative correlation between HbO_2 and HbR changes related to neurovascular coupling. Under the framework of HMS analysis, the relationship between HbO_2 and HbR in the global component is assumed to be positive. The functional and system components are separated by an empirical procedure determining the respective coefficients [38].

The HbO_2 and HbR signals and their spectra processed by the three algorithms are shown in Fig. 2. From the comparison between Fig. 2(c) and Fig. 2(f), we can see that after the three processing methods, the frequency spectrum of the HbO_2 signal decreases obviously in the full frequency band from 0.01-2 Hz, while the spectrum of the HbR signal changes little. This confirms that the HbR signal is less likely to be disturbed by scalp and global noise. For the sake of convenience, the data that are not processed by these three algorithms are referred to as raw data in the following.

2.3. Continuous wavelet transform and brain network analysis

The overall CWT and brain network analysis process is shown in Fig. 1(b). As already mentioned, CWT is able to transform the fNIRS signal from the time domain to the time-frequency domain. Based on the Morlet wavelet transform can well balance the localization of time and frequency, it is chosen as the mother wavelet function with a centre frequency of 1 [54,55]. Twenty-five logarithmic scales were selected for wavelet transformation in the frequency domain, and sixteen 30-s time windows were selected in the time domain for analysis of the wavelet transformation results. Thus, for each channel of a single subject, a 25×16 wavelet amplitude (WA) matrix was obtained. WA can be interpreted as a metric of energy to quantify the intensity of cortical activity. For every two channels, a 25×16 wavelet phase coherence (WPCO) matrix was obtained by calculating the stable level of the instantaneous phase difference between the two signals at a certain frequency during the measurement (Fig. 1(b)). The WPCO ranges from 0 to 1, and a larger value indicates that the phase difference between the two time series is more coherent at

Biomedical Optics EXPRESS

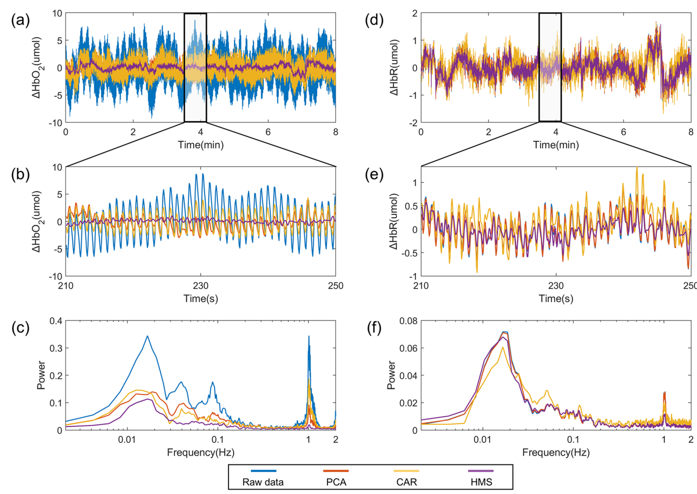

Fig. 2. Time series and spectrum of haemoglobin signal after processing by three methods of removing scalp and global blood flow noise. (a-c) Time series and spectrum of HbO2. (d-f) Time series and spectrum of HbR.

that frequency scale; therefore, it can be adopted as a measure of brain functional connectivity [56].

For each frequency scale and each time window, we can obtain the functional connectivity network with the WPCO value as the connectivity strength. Considering the need to balance false-positive and false-negative connections for constructing brain networks, the sparsity of brain networks was selected as 50% with reference to the studies of Marcel et al. and Wang et al. [57,58] That is, a WPCO value higher than the median of the connectivity matrix (24×24) was retained, a WPCO value lower than the median was set to 0, and the weight of a connection was defined as the reciprocal of the WPCO value. As a result, we complete the construction of the weight network for each time-frequency scale.

Based on the obtained weight brain network, we focused on exploring three commonly used global network topology metrics: average shortest path length, global efficiency, and weighted clustering coefficient [59], as well as three commonly used node metrics: degree, betweenness, and local efficiency [26]. These metrics, which are not necessarily correlated with each other, reflect different characteristics of brain networks. For a comprehensive description and equation of the metrics, please refer to the research of Rubinov and Sporns. All network metrics were calculated using the open-source MATLAB toolbox [60].

2.4. Test-retest reliability analysis

We used ICC (one-way random effects) to quantify the reliability of WA, WPCO and brain network metrics. In this work, we focused more on the reproducibility of the mean of repeated measurements, so the ICC was calculated as follows:

$$ICC = \frac{MS_b - MS_w}{MS_b + (k-1)MS_w} \tag{1}$$

where MS_b indicates between-subject variation, MS_w indicates within-subject variation, $k = 2$ represents the number of sessions [61,62]. In general, ICC values ranged from 0 to 1 and were classified into five test-retest reliability levels: poor (ICC < 0.2), fair (0.2 < ICC < 0.39), moderate (0.4 < ICC < 0.59), high (0.6 < ICC < 0.79), and excellent (ICC > 0.8). However, negative ICC

Biomedical Optics EXPRESS

values may occur during the calculation, the reason for which is not clear, and negative ICC values were corrected to 0 in this study [17,26,40,63].

After calculating the ICC matrix (25×16) of each time-frequency scale, the average ICC values of each cortex area were calculated for WA and nodal network metrics, and the average ICC values between each pair of brain regions were calculated for WPCO. Then, we comprehensively investigated the resting-state fNIRS scanning test-retest reliability of stroke patients in the following aspects.:

(1) Selection of frequency band and scanning duration

For the ICC matrix corresponding to each parameter of interest, the variation characteristics with frequency were observed, and the frequency band range that needs to be divided in further analysis was determined. The scanning duration was classified into eight 1-minute time windows. Two-way ANOVA was used to determine whether frequency and time have interactive effects on ICC, and the post hoc test was used to examine the effect of the choice of the analysed frequency band and the scan duration on ICC.

(2) Single channel level and brain region level

After determining the frequency band of interest, the test-retest reliability at the single channel level and brain region level was compared within each band. The single channel level refers to the average ICC value of all channels in the brain region after calculating the ICC value of a single channel, and the brain region level refers to the ICC value calculated after calculating the average measurement of brain activity or network in that brain region. The difference between them was compared by paired sample t tests in the time domain.

(3) Different preprocessing methods

In both of the above two investigating aspects, paired sample t tests were used to compare the change in ICC values corresponding to the PCA, CAR, and HMS methods relative to the ICC values of raw data.

Bonferroni correction was adopted in statistical analysis. However, the correction method is too strict, so we used different correction thresholds for the statistical analysis of different fNIRS metrics. For the T-test involved in WA, WPCO and local network metrics, $p < 0.001$ (0.05/50); For the T-test involved in global network metrics, $p < 0.0083$ (0.05/6).

3. Results

3.1. Influence of frequency band and scanning duration

The ICC matrices corresponding to the global network metric, the average WA and average local network metrics for brain regions, and the WPCO metric between brain regions were examined for all preprocessing methods. We found that the ICC values in the low-frequency band (0.01-0.08 Hz) were all smaller than those in the high-frequency band (0.145-2 Hz). The low-frequency fluctuations (0.01-0.08 Hz) of the resting state fNIRS signal are thought to possibly reflect spontaneous neural activity [15,64], 0.145-2 Hz was referred to previous studies on the division of respiratory (0.145-0.6 Hz) and heart rate activity (0.6-2 Hz) frequency bands [18,65]. Thus, we dividing the frequency scale into the two bands, which also avoided considering the influence of Mayer waves (~0.1 Hz) [49]. A two-way ANOVA was performed by combining the ICC values in the two frequency bands with eight 1-minute time windows, and it was found that the ICC of the high-frequency band was significantly higher than that of the low-frequency band for almost all the measures of interest. Additionally, the ICC values showed a tendency to increase with time, and there was a significant difference in test-retest reliability for scan durations above 4 minutes relative to those within 4 minutes. Figure 3 shows the results for several parameters across different preprocessing methods or brain regions. For a more comprehensive statistical analysis of ICC in the time and frequency domains, please refer to Supplementary File S1.

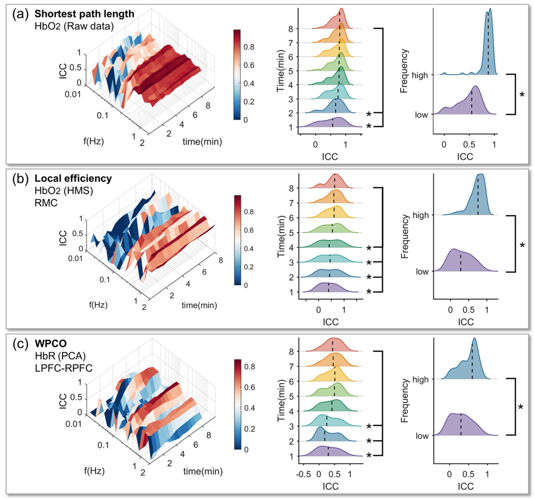

Fig. 3. Examples of ICC changes in parameters in the time and frequency domains. (a) The ICC values for short path length. (b)The mean ICC values for local efficiency of all channels (4 channels) in RMC. (c) The mean ICC values of all channel pairs (16 channel pairs) for WPCO between LPFC and RPFC. The curves of the middle column were the probability density distribution curves of ICC value per minute, the curves of the right column were the probability density distribution curves of high frequency band and low frequency band, and the area under each curve was 1. On the time scale, the ICC values of the parameters reached a relatively stable level at 8 minutes; thus, only significant results were shown compared with the ICC at 8 minutes. * indicates $p < 0.05$, high (0.145-2 Hz), low (0.01-0.08 Hz).

3.2. Comparison between single channel level and brain region level

Figures 4–8 show the ICC values corresponding to the WA, WPCO and local network metrics. The frequent appearance of four colours (corresponding to the signals of four processing modes) with slanted upwards thin arrows in the Figures implied that the ICC values were significantly larger at the brain region level than at the single channel level ($p < 0.001$, Bonferroni corrected, alpha = 0.05/50), and this is true for all the investigated metrics of interest.

3.3. Influence of preprocessing methods

The effect of the preprocessing methods on the test-retest reliability showed frequency specificity. As shown in Figs. 4–9, for all cortical activity intensity and network measures, the three preprocessing methods significantly reduced the ICC values in the high-frequency band. Generally, the ICC value of the low-frequency band was significantly improved by the three preprocessing methods, especially in local efficiency ($p < 0.001$, Bonferroni corrected, alpha = 0.05/50) and global network metrics ($p < 0.0083$, Bonferroni corrected, alpha = 0.05/6).

3.4. Comparison between single-frequency and frequency-band analysis at the low frequency level

In the usual sense, fNIRS studies are interested in the low-frequency band; therefore, the following describes the level of reliability (region level) for each metric in the low-frequency band. The single frequency level refers to the average ICC value of all scales in the low or high frequency band after calculating the ICC value of a single scale. The frequency band level refers to the ICC value calculated after calculating the average measurement of brain activity or network in low or high frequency band.

Biomedical Optics EXPRESS

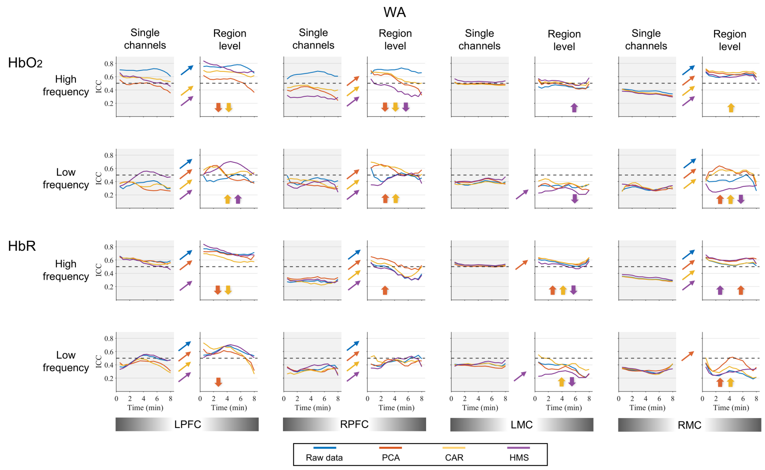

Fig. 4. Comprehensive evaluation of ICC values of 4 brain regions calculated by WA. The top two lines represent HbO2, and the bottom two lines represent HbR. The grey filling of the image refers to the single channel level, and the white filling represents the brain region level. The dotted lines in the subfigures represent ICC=0.5. The oblique upwards thin arrows of different colours in the picture mean that the ICC at the brain region level is significantly higher than that at the single channel level under different preprocessing methods (paired sample t test, $p < 0.001$, Bonferroni corrected, alpha = 0.05/6). The thick upwards or downwards arrows of different colours indicated that the three preprocessing methods significantly increased or decreased the ICC value of raw data.

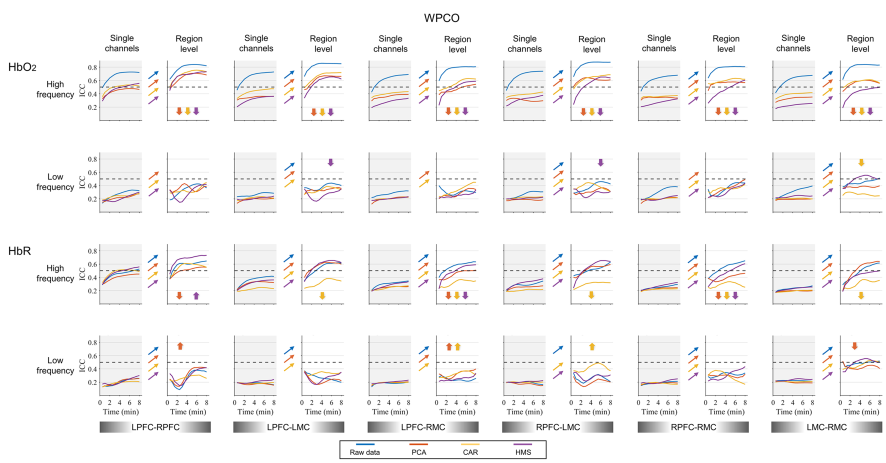

Fig. 5. Comprehensive evaluation of ICC values of 6 brain region pairs calculated by WPCO. The meaning of the subfigures and symbols can be found in the accompanying notes to Fig. 4.

(1) Single frequency level

Figure 9 demonstrates that the global network metric was able to achieve moderate to high test-retest reliability ($0.4 < ICC < 0.8$) at single frequency levels in the low-frequency band with a scan time of more than 4 minutes. In contrast, as shown in Figs. 4–8, the ICCs of the WA, WPCO and local brain network metrics in the low-frequency band exhibited only moderate, fair, or poor reliability levels ($ICC < 0.6$). In particular, the degree and betweenness show only fair or

Biomedical Optics EXPRESS

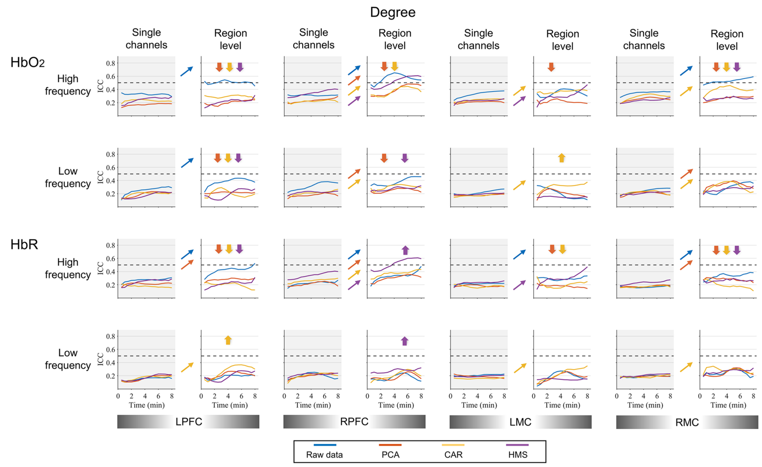

Fig. 6. Comprehensive evaluation of ICC values of 4 brain regions calculated by degree. The meaning of the subfigures and symbols can be found in the accompanying notes to Fig. 4.

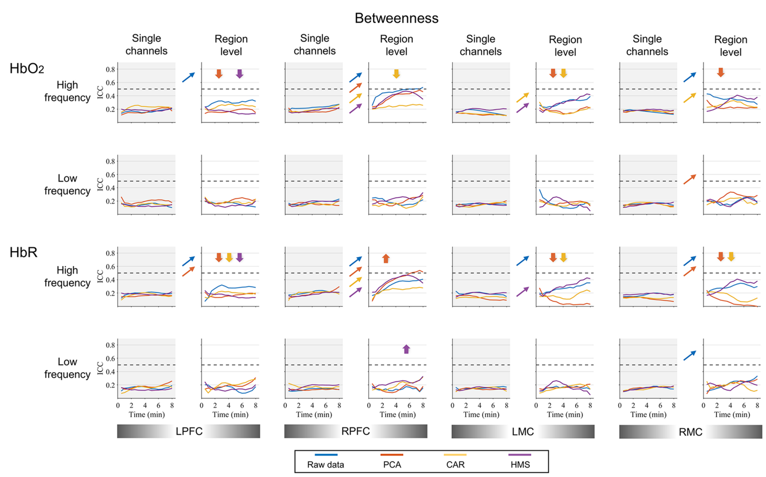

Fig. 7. Comprehensive evaluation of ICC values of 4 brain regions calculated by betweenness. The meaning of the subfigures and symbols can be found in the accompanying notes to Fig. 4.

even poor reliability levels (ICC < 0.4). The above results were performed at the level of a single frequency, which means that in the reliability evaluation of the low-frequency band, the ICC calculated at a single frequency scale within the frequency band was statistically analysed.

(2) Frequency band level

The actual fNIRS dataset analysis is concerned with the variation in haemoglobin concentration across the frequency band. In view of the unsatisfactory test-retest reliability at the single frequency level, we speculate that the analysis of the band level will improve the reliability of the fNIRS measure and supplement the ICC statistics at the band level in the low-frequency band ($p < 0.001$). For WA, ICC values improved to 0.4-0.9 (Fig. 10(a)); ICC values for local efficiency and WPCO reached 0.5-0.9 after 4 minutes (for raw data and HMS, see Fig. 10(d) and Fig. 11). The reliability of the global network parameters also reached moderate to excellent levels after 4 minutes

Biomedical Optics EXPRESS

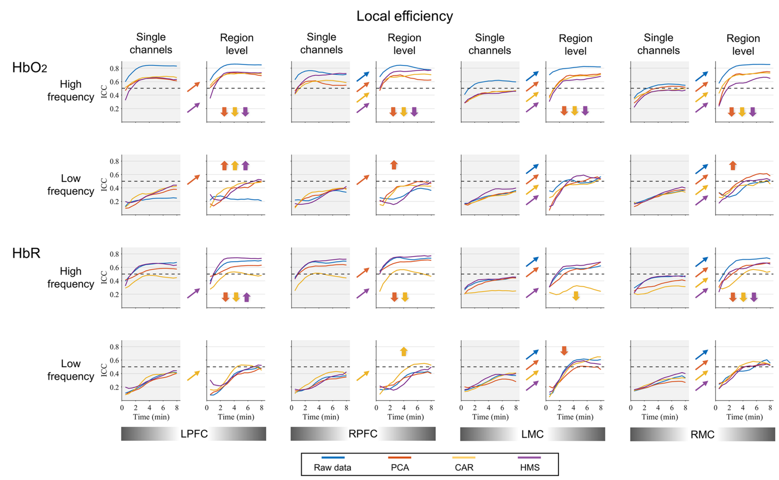

Fig. 8. Comprehensive evaluation of ICC values of 4 brain regions calculated by local efficiency. The meaning of the subfigures and symbols can be found in the accompanying notes to Fig. 4.

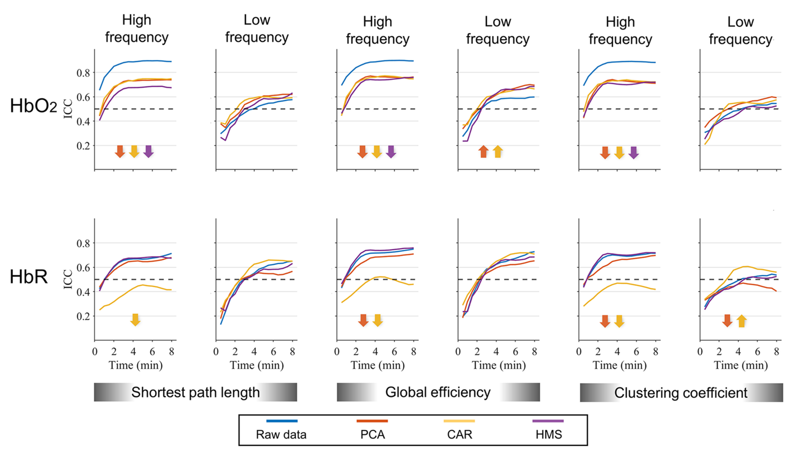

Fig. 9. Comprehensive evaluation of ICC values calculated by global network metrics. The thick upwards or downwards arrows of different colours indicated that the three preprocessing methods significantly changed the ICC value of raw data (paired sample t test, $p < 0.0083$ (Bonferroni corrected, alpha = 0.05/6).

($0.5 < ICC < 0.9$, except for the clustering coefficients of PCA-treated HbR). Although the ICC values for degree and betweenness showed an increase, they presented rather unstable reliability levels over the full time period in each brain region. Unlike the single frequency level, the three preprocessing methods reduced the test-retest reliability of degree (except for the HbR of RPFC treated by HMS), as shown in Fig. 10(b) and Fig. 10(c).

The three preprocessing methods had different effects on the test-retest reliability at the single frequency level and the frequency band level. At the single frequency level, the three preprocessing steps improved the ICC values compared to raw data overall, since PCA and HMS reduced the low-frequency ICC values of WA, WPCO and degree in some brain regions, while CAR only reduced the WPCO of LMC-RMC and the degree of LPFC. Therefore, we concluded that the CAR algorithm performs better in improving the test-retest reliability. At

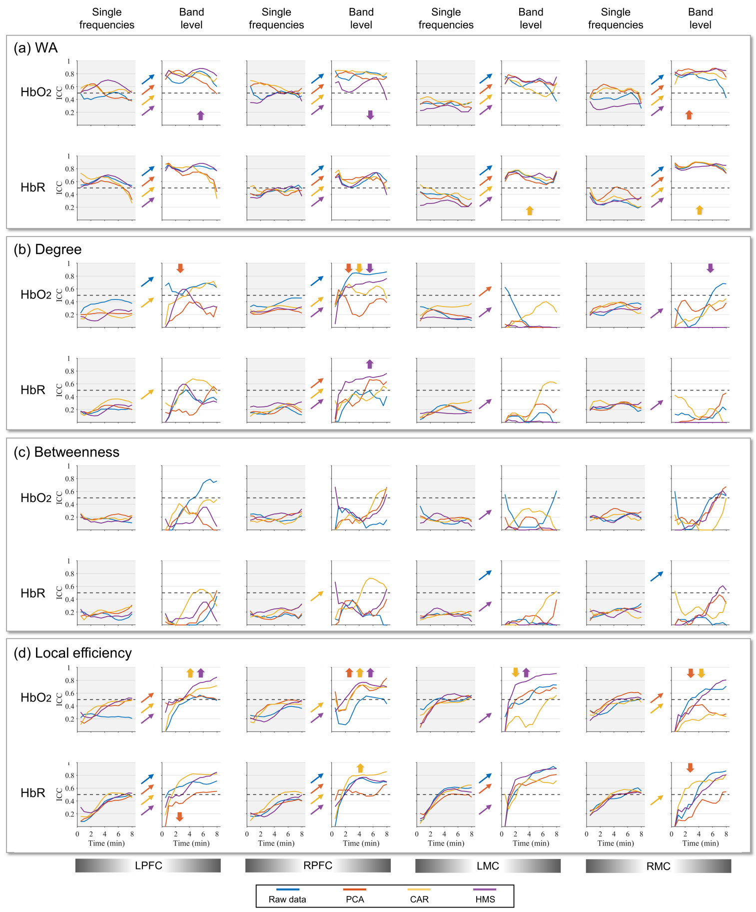

Fig. 10. Comparison of test-retest reliability of WA, degree, betweenness and local efficiency at the single frequency level and frequency band level (all at the brain region level). The grey filling of the image refers to the single frequency level, and the white filling represents the frequency band level. The dotted lines in the subfigures represent ICC = 0.5. The oblique upwards thin arrows of different colours in the picture mean that the ICC at the frequency band level is significantly higher than that at the single frequency level under different preprocessing methods (paired sample t test, $p < 0.001$). The thick upwards or downwards arrows of different colours indicated that the three preprocessing methods significantly increased or decreased the ICC value of raw data.

the frequency band level, the effect of HMS on improving ICC values is more stable across the metrics, especially for the local efficiency and global network metrics. Although HMS significantly reduced the ICC values of WA and degree for RPFC and WPCO for LPFC-LMC, the reduced ICC values were within the acceptable range (ICC values between 0.4 and 0.8 after 4 minutes).

Biomedical Optics EXPRESS

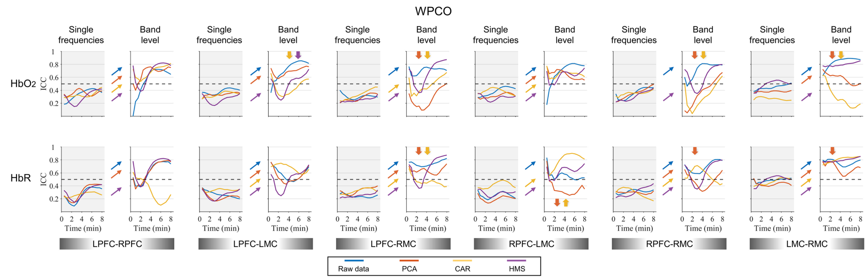

Fig. 11. Comparison of test-retest reliability of WPCO at single frequency level and frequency band level (all at the brain region level). The meaning of the subfigures and symbols can be found in the accompanying notes to Fig. 10.

4. Discussion

In this paper, we investigated the test-retest reliability of resting-state cortical activity and functional network metrics measured by fNIRS in the time-frequency domain (0.01-2 Hz, 0-8 min) for stroke patients by continuous wavelet transform and evaluated the impact on reliability of three preprocessing methods to automatically reduce system interference, PCA, CAR and HMS. Based on two fNIRS scanning sessions spaced 24 hours apart, we found that (1) the high-frequency band with higher global noise content has higher test-retest reliability than the low-frequency band, (2) the low-frequency bands require more than 4 minutes of scanning duration to ensure above-moderate reliability, (3) the fNIRS metrics are more reliable when averaged within brain area channels than at the single channel level, (4) the three automated algorithms to remove global noise have opposite effects on ICC values in high and low-frequency bands, and the HMS algorithm performs best in improving the reliability of measures of interest in low-frequency bands, (5) the test-retest reliability of degree and betweenness is fair or even poor, and the local efficiency and global network metrics reach high and excellent levels after 4 minutes.

4.1. Selection of frequency, scanning duration and analysis level

The difference in test-retest reliability between high- and low-frequency bands is an interesting finding that has not been reported in previous fNIRS studies. Studies have demonstrated high reliability of task-related activation in the frequency bands of 0.018-3 Hz [66], 0.02-0.7 Hz [22,67], 0.01-0.5 Hz [27] and resting-state brain networks in the frequency band of 0.009-0.08 Hz [25,26] without considering the effect of the choice of the analysed frequency band on the results. A previous fMRI study showed that ICC values using bandpass filtering at 0.01-0.1 Hz are significantly lower than data using bandpass filtering with wider frequency bands [17]. This may be because the higher frequency band contains more scalp and global blood flow effects, which are relatively stable in scanning stroke subjects at 1-day intervals. In contrast, the low-frequency band signal is thought to be primarily associated with local neurogenic activity, and it is difficult to ensure the constancy of the subject's thoughts during resting-state measurements, especially when disturbed by illness, resulting in relatively low anterior-posterior agreement.

From the perspective of scanning time, the results of time-frequency two-way ANOVA (Fig. 3, Supplementary File S1) and the ICC value at the frequency band level (Fig. 10- Fig. 12) show that reliability increases with scan duration and reaches above moderate levels only after 4 minutes or more (ICC > 0.4). Referring to studies of fMRI in healthy subjects, nodal degrees greater than 5 minutes were more reliable than those less than 5 minutes [17], the test-retest reliability of regional homogeneity was above the 50% level after 4 minutes [40], and longer scanning durations led to more significant reliability improvements [68]. We therefore recommend that

Biomedical Optics EXPRESS

resting-state fNIRS scans in stroke patients in clinical longitudinal studies need to last at least 4 minutes. We speculate that this observation is equally applicable in studies of healthy subjects.

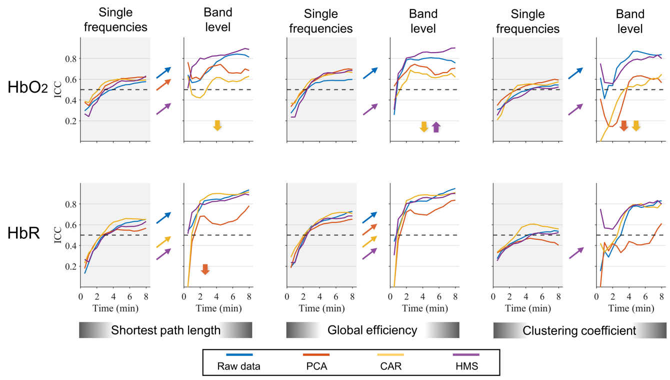

Fig. 12. Comparison of test-retest reliability of global network metrics at the single frequency level and frequency band level (all at the brain region level). The grey filling of the image refers to the single frequency level, and the white filling represents the frequency band level. The dotted lines in the subfigures represent ICC = 5. The thick upwards or downwards arrows of different colours indicated that the three preprocessing methods significantly changed the ICC value of raw data (paired sample t test, $p < 0.0083$ (Bonferroni corrected, alpha = 0.05/6).

When examining test-retest reliability, metrics at the brain region level are higher than those at the individual channel level, which is a consistent finding in fNIRS studies [23,24,69], but previous studies have been based on task activation performance, and our results suggest that this finding is also applicable to measurements of fNIRS resting cortical activity, functional connectivity and local network metrics. In addition, with the help of wavelet transform analysis, our study compared for the first time the test-retest reliability at the single frequency level and the band level. Similar to the results for the comparison of channels and brain regions, the band level was more reliable than the single frequency level. These results all seem to indicate that the larger the scale of the analysis – whether in terms of time, space or frequency – the higher the test-retest reliability tends to be.

4.2. Effect of different preprocessing methods on the test-retest reliability of each fNIRS metric

In the high frequency band, the three preprocessing methods to reduce the system noise all caused the reduction of the frequency spectrum and test-retest reliability of HBO_2 and HBR signals. This suggests that either based on the hypothesis that spatially global disturbances are consistent (PCA and CAR) [53,70] or the opposite hypothesis of functional activity-related HbO_2 and HbR (HMS) [38], it is effective to remove the interference of scalp and global blood flow. The decrease in test-retest reliability implies that the removal of the global interference step reduced within-subject variation while reducing between-subject variation more [71]. The reduction in retest reliability implies that the global interference removal step reduces intra-subject variability while reducing inter-subject variability more.

Given the higher reliability of fNIRS metrics at the brain region and frequency band levels, this paper focuses on the effect of preprocessing methods on the test-retest reliability at these two levels for low-frequency bands (Fig. 10-Fig. 12). Consistent with resting-state fMRI reports, different preprocessing methods had inconsistent effects on the test-retest reliability of different measures, even in different cortical regions of the same measure [17,63]. The performance

Biomedical Optics EXPRESS

of PCA and CAR is quite similar, which may be attributed to the high similarity between the PCA component with the greatest contribution to variance and the average time series of the whole channel [36,72], which can also be supported by the similarity of haemoglobin time series and spectrum after the two processing steps (Fig. 2). Our results show that HMS has the most consistent improvement in reliability for the fNIRS metric in the low-frequency band, especially for the local efficiency and the global network metrics. The reduction in test-retest reliability of WPCO due to the three methods was in line with our expectation. Low-frequency HbO_2 and HbR also contain the effects of scalp and global interference (such as Mayer waves) [49], which can lead to an overestimation of the functional connectivity strength. The correction techniques for non-functional components will undoubtedly weaken this correlation and reduce the test-retest reliability of functional connectivity [71].

The WA test-retest reliability of HbO_2 and HbR was at a stable high level within 0-8 minutes ($0.4 < ICC < 0.9$, see Fig. 10(a)). This shows that the measurement of resting cortical activity in stroke patients at the brain region level is reliable. Although there is a lack of research and understanding of fNIRS resting-state cortical activity in stroke patients, fMRI studies have reported abnormal low-frequency fluctuations in the parietal cortex in stroke patients compared with healthy controls [73], and the dynamic low-frequency fluctuations in auxiliary motor areas are significantly related to the improvement of motor function [74]. Clinical follow-up EEG studies have also identified spectral differences between the affected and healthy hemispheres, and the variability decreases with the rehabilitation process. The identification of targets using EEG spectral abnormalities has been reported to enhance the effectiveness of speech therapy [75–77]. In view of this, the good test-retest reliability of WA makes it possible to become a new biomarker in future stroke rehabilitation studies.

In the local network metrics, degree and betweenness show unstable reliability, and the ICC value mostly lies below 50% (Fig. 10(a) and Fig. 10(b)). Poor test-retest reliability of degree centrality [63] and betweenness [78] in network metrics in healthy subjects has also been reported previously and should therefore be used with caution in stroke clinical studies as well.

4.3. Limitations and future directions

Several limitations of current study need to be considered. The classification of high and low frequency bands in this study was first based on the investigation of the statistical results. We observed the results of ICC statistical analysis of all fNIRS metrics and found significant differences in the low frequency band (<0.1 Hz) and the high frequency band (>0.1 Hz). The low-frequency fluctuations (0.01-0.08 Hz) of the resting state fNIRS signal are thought to possibly reflect spontaneous neural activity and are also a band of interest for fMRI studies [15,64], and the low frequency band connectivity features have great potential as biomarkers for clinical applications [79]. Besides, the effect of Mayer wave interference (~ 0.1 Hz) can reduce the accuracy of fNIRS estimation and analysis [49]. Thus, we refer to previous studies on the division of respiratory (0.145-0.6 Hz) and cardiac activity (0.6-2 Hz) frequency bands [65,80], and to avoid Mayer wave interference, the low frequency range was divided into 0.01-0.08 Hz, and the high frequency range was divided into 0.145-2 Hz. Although the heartbeat and respiratory components are often removed as interferers in the fNIRS study, one of our studies demonstrated the compensation of cortical functional connectivity in the cardiac activity frequency band in stroke patients [18]. Nevertheless, the ignored 0.08-0.145 Hz was considered to cover information related to the oscillation of smooth muscles of vessels [81]. As a result, the test-retest reliability analysis of fNIRS index in this study is not comprehensive on the frequency scale.

Although PCA has been a widely used global correction method, the criteria for removing components, such as the first principal component [36,82], the first two principal components [83] and the components with 80% variance contribution [28,29], are still not uniform. We selected the last one, while different removal criteria may lead to different retest reliability. In

Biomedical Optics EXPRESS

addition, the choice of sparsity has been reported to have an impact on the value and retest reliability of network metrics [25,59,84], and it has also been demonstrated that similar statistical results were obtained under different sparsity or threshold choices [58,85]. The sparsity selection of 50% was determined by the comprehensive consideration of previous studies, nevertheless, the reporting of single sparsity may bias the test-retest reliability values.

This paper only focuses on the test-retest reliability of the resting state, while fNIRS has been used to measure the cortical responses of stroke patients under the paradigm of motor, cognitive and magnetoelectric stimulation. Future research should be extended to specific rehabilitation training paradigms. Besides, although test-retest reliability is not the only parameter to be considered when selecting an analysis strategy, it is a key variable in such an important choice [86]. If a preprocessing method does not guarantee acceptable test-retest reliability while removing interference, the robustness of the study is bound to be weakened. More work on the reliability validation of preprocessing or analysis methods needs to be done.

5. Conclusions

To our knowledge, this study provides the first comprehensive assessment of the reliability of fNIRS-based cortical activity intensity and network metrics for use in resting-state studies of stroke. We found that most fNIRS metrics achieved high to excellent levels of reliability at scanning durations of 4 minutes or more. Based on the results, we recommend that in fNIRS studies of clinical rehabilitation, (1) at least 4 minutes of resting state duration should be guaranteed to ensure the reliability of cortical activity intensity, functional connectivity, local efficiency, and global network metrics; (2) the analysis and report should be based on the average value of the channels in the brain functional area; (3) the HMS algorithm should be considered for global noise correction in the absence of short distance channels; and (4) the degree and betweenness metrics should be used with caution.

Funding. National Natural Science Foundation of China (32271370); National Key Research and Development Program of China (2020YFC2004200); Fundamental Research Funds for Central Public Welfare Research Institutes (118009001000160001).

Acknowledgements. The author would like to thank all stroke subjects who participated in the fNIRS scanning for their patient cooperation.

Disclosures. Daifa Wang declare no potential non-financial competing interests. The other authors declare no potential conflict of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the corresponding author upon reasonable request and under a licensing agreement.

Supplemental document. See Supplement 1 for supporting content.

References

1. T. Li, H. Gong, and Q. Luo, "Visualization of light propagation in visible Chinese human head for functional near-infrared spectroscopy," J. Biomed. Opt. **16**(04), 045001 (2011).
2. T. Li, H. U. I. Gong, and Q. Luo, "Mcvm: Monte Carlo modeling of photon migration in voxelized media," J. Innovative Opt. Health Sci. **03**(02), 91–102 (2010).
3. H. Ayaz, W. B. Baker, G. Blaney, D. A. Boas, H. Bortfeld, K. Brady, J. Brake, S. Brigadoi, E. M. Buckley, and S. A. Carp, "Optical imaging and spectroscopy for the study of the human brain: status report," Neurophotonics **9**(S2), S24001 (2022).
4. F. Scholkmann, S. Kleiser, A. J. Metz, R. Zimmermann, J. Mata Pavia, U. Wolf, and M. Wolf, "A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology," NeuroImage **85**(Pt 1), 6–27 (2014).
5. M. Wolf, M. Ferrari, and V. Quaresima, "Progress of near-infrared spectroscopy and topography for brain and muscle clinical applications," J. Biomed. Opt. **12**(6), 062104 (2007).
6. M. D. Pfeifer, F. Scholkmann, and R. Labruyère, "Signal processing in functional near-infrared spectroscopy (fNIRS): methodological differences lead to different statistical results," Front. Hum. Neurosci. **11**, 1 (2018).
7. K. M. Arun, K. A. Smitha, P. N. Sylaja, and C. Kesavadas, "Identifying resting-state functional connectivity changes in the motor cortex Using fNIRS During Recovery from Stroke," Brain Topogr **33**(6), 710–719 (2020).

Biomedical Optics EXPRESS

8. M. Delorme, G. Vergotte, S. Perrey, J. Froger, and I. Laffont, "Time course of sensorimotor cortex reorganization during upper extremity task accompanying motor recovery early after stroke: An fNIRS study," Restor. Neurol. Neurosci. **37**(3), 207–218 (2019).
9. P. Y. Lin, J. J. Chen, and S. I. Lin, "The cortical control of cycling exercise in stroke patients: an fNIRS study," Hum. Brain Mapp. **34**(10), 2381–2390 (2013).
10. J. S. Siegel, B. A. Seitzman, L. E. Ramsey, M. Ortega, E. M. Gordon, N. U. F. Dosenbach, S. E. Petersen, G. L. Shulman, and M. Corbetta, "Re-emergence of modular brain networks in stroke recovery," Cortex **101**, 44–59 (2018).
11. A. G. Guggisberg, P. J. Koch, F. C. Hummel, and C. M. Buetefisch, "Brain networks and their relevance for stroke rehabilitation," Clin. Neurophysiol. **130**(7), 1098–1124 (2019).
12. C. Huo, G. Xu, Z. Li, Z. Lv, Q. Liu, W. Li, H. Ma, D. Wang, and Y. Fan, "Limb linkage rehabilitation training-related changes in cortical activation and effective connectivity after stroke: a functional near-infrared spectroscopy study," Sci. Rep. **9**(1), 6226 (2019).
13. I. Miyai, H. Yagura, I. Oda, I. Konishi, H. Eda, T. Suzuki, and K. J. A. O. N. Kubota, "Premotor cortex is involved in restoration of gait in stroke," Ann. Neurol. **52**, 188–194 (2002).
14. H. Fujimoto, M. Mihara, N. Hattori, M. Hatakenaka, T. Kawano, H. Yagura, I. Miyai, and H. J. N. Mochizuki, "Cortical changes underlying balance recovery in patients with hemiplegic stroke," NeuroImage **85**, 547–554 (2014).
15. G. Xu, C. Huo, J. Yin, W. Li, H. Xie, X. Li, Z. Li, Y. Wang, and D. Wang, "Effective brain network analysis in unilateral and bilateral upper limb exercise training in subjects with stroke," Med. Phys. **49**(5), 3333–3346 (2022).
16. C. A. Lawler, I. M. Wiggins, R. S. Dewey, and D. E. J. C. I. I. Hartley, "The use of functional near-infrared spectroscopy for measuring cortical reorganisation in cochlear implant users: a possible predictor of variable speech outcomes?" Cochlear Implants Int. **16**, S30–S32 (2015).
17. M. Andellini, V. Cannata, S. Gazzellini, B. Bernardi, and A. Napolitano, "Test-retest reliability of graph metrics of resting state MRI functional brain networks: A review," J. Neurosci. Methods **253**, 183–192 (2015).
18. Q. Tan, M. Zhang, Y. Wang, M. Zhang, Y. Wang, Q. Xin, B. Wang, and Z. Li, "Frequency-specific functional connectivity revealed by wavelet-based coherence analysis in elderly subjects with cerebral infarction using NIRS method," Med. Phys. **42**(9), 5391–5403 (2015).
19. M. Bonstrup, R. Schulz, G. Schon, B. Cheng, J. Feldheim, G. Thomalla, and C. Gerloff, "Parietofrontal network upregulation after motor stroke," NeuroImage Clin. **18**, 720–729 (2018).
20. A. R. Carter, G. L. Shulman, and M. Corbetta, "Why use a connectivity-based approach to study stroke and recovery of function?" NeuroImage **62**(4), 2271–2280 (2012).
21. C. Grefkes and G. R. Fink, "Connectivity-based approaches in stroke and recovery of function," Lancet Neurol. **13**(2), 206–216 (2014).
22. M. M. Plichta, M. J. Herrmann, C. G. Baehne, A. C. Ehlis, M. M. Richter, P. Pauli, and A. J. Fallgatter, "Event-related functional near-infrared spectroscopy (fNIRS): are the measurements reliable?" NeuroImage **31**(1), 116–124 (2006).
23. T. Kono, K. Matsuo, K. Tsunashima, K. Kasai, R. Takizawa, M. A. Rogers, H. Yamasue, T. Yano, Y. Taketani, and N. Kato, "Multiple-time replicability of near-infrared spectroscopy recording during prefrontal activation task in healthy men," Neurosci. Res. **57**(4), 504–512 (2007).
24. A. Blasi, S. Lloyd-Fox, M. H. Johnson, and C. Elwell, "Test-retest reliability of functional near infrared spectroscopy in infants," Neurophotonics **1**(2), 025005 (2014).
25. H. Niu, Z. Li, X. Liao, J. Wang, T. Zhao, N. Shu, X. Zhao, and Y. He, "Test-retest reliability of graph metrics in functional brain networks: a resting-state fNIRS study," PLoS One **8**(9), e72425 (2013).
26. M. Wang, Z. Yuan, and H. Niu, "Reliability evaluation on weighted graph metrics of fNIRS brain networks," Quant. Imaging Med. Surg **9**(5), 832–841 (2019).
27. I. M. Wiggins, C. A. Anderson, P. T. Kitterick, and D. E. Hartley, "Speech-evoked activation in adult temporal cortex measured using functional near-infrared spectroscopy (fNIRS): Are the measurements reliable?" Hear. Res. **339**, 142–154 (2016).
28. H. Santosa, X. Zhai, F. Fishburn, P. J. Sparto, and T. J. Huppert, "Quantitative comparison of correction techniques for removing systemic physiological signal in functional near-infrared spectroscopy studies," Neurophotonics **7**(03), 035009 (2020).
29. A. Bonilauri, F. Sangiuliano Intra, G. Baselli, and F. Baglio, "Assessment of fNIRS signal processing pipelines: towards clinical applications," Appl. Sci. **12**(1), 316 (2021).
30. T. Li, Y. Li, Y. Sun, M. Duan, and L. Peng, "Effect of head model on Monte Carlo modeling of spatial sensitivity distribution for functional near-infrared spectroscopy," J. Innovative Opt. Health Sci. **08**(05), 1550024 (2015).
31. X. Fang, B. Pan, W. Liu, Z. Wang, and T. Li, "Effect of scalp hair follicles on NIRS quantification by Monte Carlo simulation and Visible Chinese Human Dataset," IEEE Photonics J. **10**(5), 1–10 (2018).
32. T. Li, C. Xue, P. Wang, Y. Li, and L. Wu, "Photon penetration depth in human brain for light stimulation and treatment: a realistic Monte Carlo simulation study," J. Innovative Opt. Health Sci. **10**(05), 1743002 (2017).
33. T. Li, Y. Lin, Y. Shang, L. He, C. Huang, M. Szabunio, and G. Yu, "Simultaneous measurement of deep tissue blood flow and oxygenation using noncontact diffuse correlation spectroscopy flow-oximeter," Sci. Rep. **3**(1), 1358 (2013).
34. B. Pan, C. Huang, X. Fang, X. Huang, and T. Li, "Noninvasive and sensitive optical assessment of brain death," J. Biophotonics **12**, e201800240 (2019).
35. H. Santosa, A. Aarabi, S. B. Perlman, and T. J. Huppert, "Characterization and correction of the false-discovery rates in resting state connectivity using functional near-infrared spectroscopy," J. Biomed. Opt. **22**(5), 055002 (2017).

Biomedical Optics EXPRESS

36. L. M. Hocke, I. K. Oni, C. C. Duszynski, A. V. Corrigan, B. D. Frederick, and J. F. Dunn, "Automated processing of fNIRS data-a visual guide to the pitfalls and consequences," Algorithms **11**(5), 67 (2018).
37. G. Bauernfeind, S. C. Wriessnegger, I. Daly, and G. R. Muller-Putz, "Separating heart and brain: on the reduction of physiological noise from multichannel functional near-infrared spectroscopy (fNIRS) signals," J. Neural Eng. **11**(5), 056010 (2014).
38. J.-C. Baron, T. Yamada, S. Umeyama, and K. Matsuda, "Separation of fNIRS signals into functional and systemic components based on differences in hemodynamic modalities," PLoS One **7**(11), e50024 (2012).
39. H. Zhang, Y. J. Zhang, C. M. Lu, S. Y. Ma, Y. F. Zang, and C. Z. Zhu, "Functional connectivity as revealed by independent component analysis of resting-state fNIRS measurements," NeuroImage **51**(3), 1150–1161 (2010).
40. X. N. Zuo, T. Xu, L. Jiang, Z. Yang, X. Y. Cao, Y. He, Y. F. Zang, F. X. Castellanos, and M. P. Milham, "Toward reliable characterization of functional homogeneity in the human brain: preprocessing, scan duration, imaging resolution and computational space," NeuroImage **65**, 374–386 (2013).
41. Eugene B. Postnikov, Elena A. Lebedeva, and Anastasia I. Lavrova, "Computational implementation of the inverse continuous wavelet transform without a requirement of the admissibility condition," Appl. Math. Comput. **282**, 128–136 (2016).
42. Z. Li, Y. Wang, Y. Li, Y. Wang, J. Li, and L. Zhang, "Wavelet analysis of cerebral oxygenation signal measured by near infrared spectroscopy in subjects with cerebral infarction," Microvasc. Res. **80**(1), 142–147 (2010).
43. A. K. Rehme, G. R. Fink, D. Y. von Cramon, and C. Grefkes, "The role of the contralesional motor cortex for motor recovery in the early days after stroke assessed with longitudinal FMRI," Cereb Cortex **21**(4), 756–768 (2011).
44. C. H. Park, W. H. Chang, S. H. Ohn, S. T. Kim, O. Y. Bang, A. Pascual-Leone, and Y. H. Kim, "Longitudinal changes of resting-state functional connectivity during motor recovery after stroke," Stroke **42**(5), 1357–1362 (2011).
45. M. A. Yucel, A. V. Luhmann, and F. Scholkmann, *et al.*, "Best practices for fNIRS publications," Neurophotonics **8**(01), 012101 (2021).
46. F. Scholkmann and M. Wolf, "General equation for the differential pathlength factor of the frontal human head depending on wavelength and age," J. Biomed. Opt. **18**(10), 105004 (2013).
47. C. M. Aasted, M. A. Yucel, R. J. Cooper, J. Dubb, D. Tsuzuki, L. Becerra, M. P. Petkov, D. Borsook, I. Dan, and D. A. Boas, "Anatomical guidance for functional near-infrared spectroscopy: AtlasViewer tutorial," Neurophotonics **2**(2), 020801 (2015).
48. X. Hou, Z. Zhang, C. Zhao, L. Duan, Y. Gong, Z. Li, and C. Zhu, "NIRS-KIT: a MATLAB toolbox for both resting-state and task fNIRS data analysis," Neurophotonics **8**(01), 010802 (2021).
49. M. A. Yucel, J. Selb, C. M. Aasted, P. Y. Lin, D. Borsook, L. Becerra, and D. A. Boas, "Mayer waves reduce the accuracy of estimated hemodynamic response functions in functional near-infrared spectroscopy," Biomed. Opt. Express **7**(8), 3078–3088 (2016).
50. X. Si, S. Xiang, L. Zhang, S. Li, K. Zhang, and D. Ming, "Acupuncture With deqi Modulates the Hemodynamic Response and Functional Connectivity of the Prefrontal-Motor Cortical Network," Front. Neurosci. **15**, 693623 (2021).
51. F. A. Fishburn, R. S. Ludlum, C. J. Vaidya, and A. V. Medvedev, "Temporal Derivative Distribution Repair (TDDR): A motion correction method for fNIRS," NeuroImage **184**, 171–179 (2019).
52. T. J. Huppert, S. G. Diamond, M. A. Franceschini, and D. A. J. A. O. Boas, "HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain," Appl. Opt. **48**(10), D280–D298 (2009).
53. G. Pfurtscheller, G. Bauernfeind, S. C. Wriessnegger, and C. Neuper, "Focal frontal (de)oxyhemoglobin responses during simple arithmetic," Int. J. Psychophysiol. **76**(3), 186–192 (2010).
54. A. Grinsted, J. C. Moore, and S. Jevrejeva, "Application of the cross wavelet transform and wavelet coherence to geophysical time series," Nonlin. Processes Geophys. **11**(5/6), 561–566 (2004).
55. G. Xu, M. Zhang, Y. Wang, Z. Liu, C. Huo, Z. Li, and M. Huo, "Functional connectivity analysis of distracted drivers based on the wavelet phase coherence of functional near-infrared spectroscopy signals," PLoS One **12**(11), e0188329 (2017).
56. T. Zhang, G. Xu, C. Huo, W. Li, Z. Li, and W. Li, "Cortical hemodynamic response and networks in children with cerebral palsy during upper limb bilateral motor training," J. Biophotonics, **16**, e202200326 (2023).
57. M. A. de Reus and M. P. van den Heuvel, "Estimating false positives and negatives in brain networks," NeuroImage **70**, 402–409 (2013).
58. L. Wang, C. Yu, H. Chen, W. Qin, Y. He, F. Fan, Y. Zhang, M. Wang, K. Li, Y. Zang, T. S. Woodward, and C. Zhu, "Dynamic functional reorganization of the motor execution network after stroke," Brain **133**(4), 1224–1238 (2010).
59. M. Drakesmith, K. Caeyenberghs, A. Dutt, G. Lewis, A. S. David, and D. K. Jones, "Overcoming the effects of false positives and threshold bias in graph theoretical analyses of neuroimaging data," NeuroImage **118**, 313–333 (2015).
60. M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," NeuroImage **52**(3), 1059–1069 (2010).
61. J. Yue, N. Zhao, Y. Qiao, Z. J. Feng, Y. S. Hu, Q. Ge, T. Q. Zhang, Z. Q. Zhang, J. Wang, and Y. F. Zang, "Higher reliability and validity of Wavelet-ALFF of resting-state fMRI: From multicenter database and application to rTMS modulation," Hum. Brain Mapp. **44**(3), 1105–1117 (2023).
62. J. Wang, Y. Ren, X. Hu, V. T. Nguyen, L. Guo, J. Han, and C. C. Guo, "Test-retest reliability of functional connectivity networks during naturalistic fMRI paradigms," Hum. Brain Mapp. **38**(4), 2226–2241 (2017).

Biomedical Optics EXPRESS

63. S. Holiga, F. Sambataro, C. Luzy, G. Greig, N. Sarkar, R. J. Renken, J. C. Marsman, S. A. Schobel, A. Bertolino, and J. Dukart, "Test-retest reliability of task-based and resting-state blood oxygen level dependence and cerebral blood flow measures," PLoS One **13**(11), e0206583 (2018).
64. C. J. Murray, T. Vos, R. Lozano, M. Naghavi, A. D. Flaxman, C. Michaud, M. Ezzati, K. Shibuya, J. A. Salomon, and S. Abdalla, "Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010," Lancet **380**(9859), 2197–2223 (2012).
65. Y. Shiogai, A. Stefanovska, and P. V. McClintock, "Nonlinear dynamics of cardiovascular ageing," Phys. Rep. **488**(2-3), 51–110 (2010).
66. M. Schecklmann, A. C. Ehlis, M. M. Plichta, and A. J. Fallgatter, "Functional near-infrared spectroscopy: a long-term reliable tool for measuring brain activity during verbal fluency," NeuroImage **43**(1), 147–155 (2008).
67. M. M. Plichta, M. J. Herrmann, C. G. Baehne, A. C. Ehlis, M. M. Richter, P. Pauli, and A. J. Fallgatter, "Event-related functional near-infrared spectroscopy (fNIRS) based on craniocerebral correlations: reproducibility of activation?" Hum. Brain Mapp. **28**(8), 733–741 (2007).
68. R. M. Birn, E. K. Molloy, R. Patriat, T. Parker, T. B. Meier, G. R. Kirk, V. A. Nair, M. E. Meyerand, and V. Prabhakaran, "The effect of scan length on the reliability of resting-state fMRI connectivity estimates," NeuroImage **83**, 550–558 (2013).
69. Y. Kakimoto, Y. Nishimura, N. Hara, M. Okada, H. Tanii, and Y. Okazaki, "Intrasubject reproducibility of prefrontal cortex activities during a verbal fluency task over two repeated sessions using multi-channel near-infrared spectroscopy," Psychiatry Clin. Neurosci. **63**(4), 491–499 (2009).
70. T. Sato, I. Nambu, K. Takeda, T. Aihara, O. Yamashita, Y. Isogaya, Y. Inoue, Y. Otaka, Y. Wada, M. Kawato, M. A. Sato, and R. Osu, "Reduction of global interference of scalp-hemodynamics in functional near-infrared spectroscopy using short distance probes," NeuroImage **141**, 120–132 (2016).
71. R. M. Birn, M. D. Cornejo, E. K. Molloy, R. Patriat, T. B. Meier, G. R. Kirk, V. A. Nair, M. E. Meyerand, and V. Prabhakaran, "The influence of physiological noise correction on test-retest reliability of resting-state functional connectivity," Brain Connect **4**(7), 511–522 (2014).
72. F. Carbonell, P. Bellec, and A. J. B. C. Shmuel, "Global and system-specific resting-state fMRI fluctuations are uncorrelated: principal component analysis reveals anti-correlated networks," Brain Connect **1**(6), 496–510 (2011).
73. J. Zhu, Y. Jin, K. Wang, Y. Zhou, Y. Feng, M. Yu, and X. Jin, "Frequency-dependent changes in the regional amplitude and synchronization of resting-state functional MRI in stroke," PLoS One **10**(4), e0123850 (2015).
74. J. Chen, D. Sun, Y. Shi, W. Jin, Y. Wang, Q. Xi, and C. Ren, "Dynamic Alterations in Spontaneous Neural Activity in Multiple Brain Networks in Subacute Stroke Patients: A Resting-State fMRI Study," Front. Neurosci. **12**, 994 (2018).
75. J. Lanzone, M. A. Colombo, S. Sarasso, F. Zappasodi, M. Rosanova, M. Massimini, V. Di Lazzaro, and G. Assenza, "EEG spectral exponent as a synthetic index for the longitudinal assessment of stroke recovery," Clin. Neurophysiol. **137**, 92–101 (2022).
76. K. Laaksonen, L. Helle, L. Parkkonen, E. Kirveskari, J. P. Makela, S. Mustanoja, T. Tatlisumak, M. Kaste, and N. Forss, "Alterations in spontaneous brain oscillations during stroke recovery," PLoS One **8**(4), e61146 (2013).
77. P. P. Shah-Basak, G. Sivaratnam, S. Teti, A. Francois-Nienaber, M. Yossofzai, S. Armstrong, S. Nayar, R. Jokel, and J. Meltzer, "High definition transcranial direct current stimulation modulates abnormal neurophysiological activity in post-stroke aphasia," Sci. Rep. **10**(1), 19625 (2020).
78. J.-H. Wang, X.-N. Zuo, S. Gohel, M. P. Milham, B. B. Biswal, and Y. J. P. O. He, "Graph theoretical analysis of functional brain networks: test-retest evaluation on short-and long-term resting-state functional MRI data," PLoS One **6**(7), e21976 (2011).
79. S. Shahdadian, X. Wang, S. Kang, C. Carter, A. Chaudhari, and H. Liu, "Prefrontal cortical connectivity and coupling of infraslow oscillation in the resting human brain: a 2-channel broadband NIRS study," Cereb Cortex Commun **3**(3), tgac033 (2022).
80. H. D. Kvernmo, A. Stefanovska, K. A. Kirkebøen, and K. Kvernebo, "Oscillations in the human cutaneous blood perfusion signal modified by endothelium-dependent and endothelium-independent vasodilators," Microvasc. Res. **57**(3), 298–309 (1999).
81. H. Obrig, M. Neufang, R. Wenzel, M. Kohl, J. Steinbrink, K. Einhaupl, and A. Villringer, "Spontaneous low frequency oscillations of cerebral hemodynamics and metabolism in human adults," NeuroImage **12**(6), 623–639 (2000).
82. S. L. Novi, R. B. Rodrigues, and R. C. Mesquita, "Resting state connectivity patterns with near-infrared spectroscopy data of the whole head," Biomed. Opt. Express **7**(7), 2524–2537 (2016).
83. B. Khan, F. Tian, K. Behbehani, M. I. Romero, M. R. Delgado, N. J. Clegg, L. Smith, D. Reid, H. Liu, and G. Alexandrakis, "Identification of abnormal motor cortex activation patterns in children with cerebral palsy by functional near-infrared spectroscopy," J. Biomed. Opt. **15**(3), 036008 (2010).
84. M. P. van den Heuvel, S. C. de Lange, A. Zalesky, C. Seguin, B. T. T. Yeo, and R. Schmidt, "Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations," NeuroImage **152**, 437–449 (2017).
85. J. D. Rudie, J. A. Brown, D. Beck-Pancer, L. M. Hernandez, E. L. Dennis, P. M. Thompson, S. Y. Bookheimer, and M. Dapretto, "Altered functional and structural brain network organization in autism," NeuroImage Clin. **2**, 79–94 (2012).
86. S. Nakagawa and H. Schielzeth, "Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists," Biological Reviews **85**(4), 935–956 (2010).